PATENT APPLICATION
Docket No. 2705-152

## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

| | | | |
|---|---|---|---|
| Inventor(s): | Luke K. Surazski, et al. | Confirmation No. | 8755 |
| Serial No. | 09/843,596 | Examiner: | Chi Ho A. LEE |
| Filed: | April 26, 2001 | Group Art Unit: | 2663 |

For:      DEVICES, METHODS AND SOFTWARE FOR GENERATING INDEXING
          METATAGS IN REAL TIME FOR A STREAM OF DIGITALLY STORED
          VOICE DATA

Date:     February 28, 2006

Assistant Commissioner for Patents
Washington, D.C. 20231

### DECLARATION OF PASCAL HUART

I, PASCAL HUART, declare the following.

1.      I am one of the co-inventors of the subject matter described in the present pending
patent application titled DEVICES, METHODS AND SOFTWARE FOR GENERATING
INDEXING METATAGS IN REAL TIME FOR A STREAM OF DIGITALLY STORED
VOICE DATA.

2.      I have first hand knowledge of the conception, the development and manufacture,
and the reduction to practice of the invention.

3.      I currently work for Cisco Technology, Inc., having a business mailing address of
170 West Tasman Drive, San Jose, CA 95134-1706.

4.  Prior to the effective date of U.S. patent no. 6,947,598 to Yogeshwar and others, I and the other named inventors of the application we reference above (i.e., U.S. patent application serial no. 09/843,596) conceived of and developed devices, software and methods for generating indexing metatags in real time for a stream of digitally stored voice data as described and claimed in the application.

5.  Attached as Exhibit A is a computerized entry establishing conception of the invention prior to the effective date (i.e., filing date) of the Yogeshwar reference. The Background and Summary subheadings in Exhibit A describe the invention as described and claimed in the application. The date of the computerized entry has been redacted out of Exhibit A, but was made prior to Yogeshwar's filing date of April 20, 2001. We have redacted other portions of Exhibit A because they do not relate to the invention.

6.  Work on the invention was conducted continuously from a date prior to the effective date, until the date of filing of the above referenced patent application, and thereafter.

I, the undersigned, declare that all statements made of my own knowledge are true, and that all statements made on information and belief are believed to be true. I understand that willful false statements are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United Stated Code. And I understand that willful false statements may jeopardize the validity of any patent issuing on the application.

DATED this __28__ day of __February__, 2006.

Pascal Declaration                          Page 2 of 2                          Do. No. 2705-152
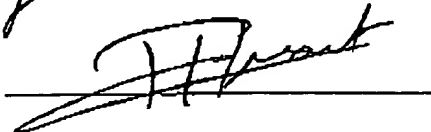
## EXHIBIT A

Edit    ☐Print View    Watch ☐    Patent Details    Review

# Devices, methods and software for generating indexing metatags in real time for a stream of digitally stored voice data

| CPOL No. 9600 : 1 | Seq No.: 3583 | Status: Pending | Submitted: ▆ | Modified : ▆ |

## Portfolio Manager

Laurie Mintz (lamintz)          **Address:** ▆          **Email:** lamintz@cisco.com
                                                        **Phone:** ▆

## Idea Details

**Inventors:** Email Inventors

Pascal Huart (phuart)     **Phone:** ▆     **Manager:** ▆     **Dept:** ▆

**Type:** Regular     **Division:** ▆     **Site:** --     **Info:** Cisco Directory

---

Michael Knappe (mknappe)          ▆

---

Luke Surazski (lsurazsk)     **Phone:** ▆     **Manager:** ▆     **Dept:** ▆

**Type:** Regular     **Division:** ▆     **Site:** ▆     **Info:** Cisco Directory

---

**Background:** A feature that is currently not implemented in our VoIP solution is the ability to record calls and conferences.  This feature is highly desirable for archiving conference calls and recording conversations that could later be used as records of a transaction, etc.  Even with compression, voice data requires a good deal of bandwidth for storage.  For this reason, when recording any type of call, it is desirable to record (store to some mass storage device) the mix of the call/conversation rather than the individual streams of the participants.

Traditionally when calls are recorded they can only be played back.  This makes scanning the call for information or a particular speaker difficult. This invention discusses how using VoIP these problems can be solved while keeping in mind the above requirement to use up as little space as possible.

**Possible Prior Art:** ---

## EXHIBIT A

**Summary** This invention proposes methods of recording calls and conference calls while facilitating scanning for specific speaker and important information.

Instance 1:
As mentioned above, when a call or conference call is recorded it is typically recorded as a single stream rather than a stream for each individual speaker. This ensures that the recorded conference does not consume a great deal of bandwidth. Typically VoIP conferencing solutions incorporate some sort of speaker selection/VAD mechanism so that only the active speakers are present in the audio stream produced by the conference bridge.

The first instance of this invention proposes that when the speaker selection algorithm detects a new active speaker  then it forwards a signaling message with the speaker ID and an RTP timestamp to the recording device. The recording device uses these signaling messages to construct a table consisting of speaker ID's and timestamps which get saved along with the audio data from the conference bridge.

When the conference is being played back, this information can be sent back as signaling messages to the playback client so that users can identify who is speaking at any point in time. Furthermore, the user could use some form of user interface to direct the playback server to skip to the next instance of user X's speech. The playback server can skip to any instance of any speaker's audio by doing a simple lookup in the table that was constructed during the recording.

When the index is constructed the RTP timestamps should be converted to a standard user-readable format such as SMPTE. This facilitates browsing of the archived call audio as RTP time-stamps may be meaningless to the average user.

Instance 2:
The second instance of the invention deals with marking the recorded audio stream for important pieces of information. This part of the invention could be used for skipping to important content during playback. For instance, it might be used by a stock broker to mark the point in the conversation with a client when the client gives an account number and then again when the client specifies the details of the trade. This instance allows the listener of the playback to skip the less significant parts of the conversation.

During the recording of a conversation, the user of this invention should be presented with a softkey (as on the Cisco 7960 or on the softphone) that triggers the sending of a signaling message to the recording device. This signaling message should contain an RTP sequence number. The recording device should store the signaling messages that it receives in a table.

During playback, through some sort of user interface on the playback client, the listener should be able to skip to the marked portions of the recorded message by sending signaling messages to the playback server. In a generic application the user might be presented with "Next Index" and "Previous Index" markers. This instance of the invention could be used for more customized application in which the user of the invention is

## EXHIBIT A

presented with specific marking keys (such as ADDRESS, ACCOUNT NUMBER, TRANSACTION INFORMATION) and the listener can search for specific pieces of data in the recorded conversation.  Specifically this sort of feature might be useful for some sort of automated process that scans recorded audio and extracts specific information from it. (ie through some speech recognition process)

**Restatement:** ---

**Advantages:** This invention uses the inherent features of voice over IP to enhance the
usability of recorded calls.

In particular it provides methods of:
 -> Scanning recorded conversations for particular speakers while
        still recording the conversation as one stream
 -> Scanning recorded conversations for important or even specific
        pieces of information.
Using a SMPTE or other user-readable time-stamping format for the index
file/archived audio facilitates scanning the recorded audio.